



## **1351.0.55.045 - Research Paper: Assessing the Quality of Linking School Enrolment Records to 2011 Census Data: Deterministic Linkage Methods, Dec 2013**

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 13/12/2013 First Issue

---

## **Summary**

### **About this Release**

Deterministic record linkage is a process of locating records pertaining to the same individual from multiple data sources. This type of linkage is most applicable where characteristics from the different sources are reported consistently, so as to uniquely identify the individual. It is less applicable in instances where there are problems with data quality, or where the reported characteristics cannot ensure the unique identification of matching records.

The accuracy of matches identified by deterministic record linkage is intrinsically high. By contrast, probabilistic record linkage is useful where characteristics may differ among data sources although they genuinely belong to the same person (e.g. the way the string of characters in a name or address is recorded). Compared with deterministic linkage, probabilistic methods can return higher numbers of true matches but this may be at the expense of higher numbers of false links as well.

This paper compares the quality of several integrated datasets created by the application of probabilistic and deterministic record linkage methods for the ABS Census Data Enhancement Education Quality Study. This study linked 2011 Census data with government school enrolment records from Queensland, South Australia, Tasmania and the Northern Territory as part of feasibility testing for using data integration to expand the evidence base for education and training policy. Four examples of deterministic record linkage are compared against (i) the benchmark of probabilistic record linkage using name and address (Gold standard), and (ii) probabilistic record linkage conducted without name and address information (Bronze standard).

A significant finding is that deterministic linking utilising the SLK581 (a Statistical Linkage Key which includes coded name information) approaches the quality of Gold standard linkage, and may provide an acceptable alternative to Gold standard linkage in instances where full name and address information is not available. In comparison to probabilistic linkage, deterministic methods for Bronze standard linkage generally achieved lower linkage rates but higher quality in terms of linkage accuracy and match-link rate.

[Download Research Paper](#)